

STIC-Biotech/ChemLib

168351

ME

From: Myers, Carla  
Sent: Wednesday, October 12, 2005 7:00 AM  
To: STIC-Biotech/ChemLib  
Subject: sequence search 09/887941

Please do an oligomer search in commercial and interference files for fragments of SEQ ID NO: **11 and 12**  
- please limit the search results to nucleic acids of a length of 50 nucleotides or less.

Please provide a printout of the first 40 results.

The CRF has been entered; [http://expoweb1:8001/cgi-bin/expo/BioInfo/bioquery.pl?APPL\\_ID=09887941](http://expoweb1:8001/cgi-bin/expo/BioInfo/bioquery.pl?APPL_ID=09887941)

Thank you

Carla Myers  
AU 1634  
Remsen Bldg / Rm 2E79  
Mailbox: REM 2C70  
571-272-0747

RECEIVED  
OCT 12 2005  
STIC/BIOTECH/CHM LIB  
(STIC)

\*\*\*\*\*

Searcher: \_\_\_\_\_  
Searcher Phone: \_\_\_\_\_  
Date Searcher Picked up: 10/12/05  
Date completed: 10/12/05  
Searcher Prep Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search  
NA# 2 AA# \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure #: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIS: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: DSH  
WWW/Internet: \_\_\_\_\_  
Other (Specify): \_\_\_\_\_

OM nucleic - nucleic search, using sw model

4393.656 Million cell updates/sec

Title: US-09-887-941B-12  
Perfect score: 100

Sequence: 1 gggcgcgcagggagttca.....ttqaqaaccqaqaacacac 196

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 34239544 segs, 19032134700 residues

Word size :

Total number of hits satisfying chosen parameters: 159776

Minimum	DB	seq	length:	0
Maximum	DB	seq	length:	1

Maximum DB seq length: 50

Post-processing: Listing first 45 summaries

Database :

```

1:  gb_est1:*
2:  gb_est2:*
3:  gb_htc:*
4:  gb_est3:*
5:  gb_est4:*
6:  gb_est5:*
7:  gb_est6:*
8:  gb_gss1:*
9:  gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	14	7.1	50	1	AU107206	AU107206 AU107206
C 2	14	7.1	50	1	AU107207	AU107207 AU107207
C 3	14	7.1	50	1	AU107210	AU107210 AU107210
C 4	14	7.1	50	1	AU107211	AU107211 AU107211
C 5	14	7.1	50	1	AU107213	AU107213 AU107213
C 6	14	7.1	50	1	AU107215	AU107215 AU107215
C 7	13	6.6	43	2	BE407584	BE407584 60129909
C 8	13	6.6	50	1	AU107216	AU107216 AU107216
C 9	12	6.1	22	8	AZ481923	AZ481923 IM030610
C 10	12	6.1	25	8	AZ654166	AZ654166 IM0528M1
C 11	12	6.1	29	8	AZ804761	AZ804761 2M0065B01
C 12	12	6.1	31	1	A1582089	A1582089 ar96c09.x
C 13	12	6.1	32	7	CO792123	CO792123 NM0148.A
C 14	12	6.1	34	9	CR360566	CR360566 Arabidops
C 15	12	6.1	39	8	AZ781791	AZ781791 2M002112
C 16	12	6.1	41	8	AZ361377	AZ361377 IM0106J07
C 17	12	6.1	41	8	BZ766262	BZ766262 SALK_1365
C 18	12	6.1	43	9	AG200058	AG200058 Pen_Frog1
C 19	12	6.1	45	4	BT754711	BT754711 603025415
C 20	12	6.1	46	8	BZ766261	BZ766261 SALK_1365
C 21	12	6.1	48	1	AA000245	AA000245 mg33509.x
C 22	12	6.1	48	8	AQ025339	AQ025339 EB(X)1179
C 23	12	6.1	50	9	CR167367	CR167367 Forward s
C 24	11	5.6	16	9	AJ589882	AJ589882 Arabidops

C	25	11	5.6	20	6	C01261	C01261	HUMGS000799
C	26	11	5.6	22	8	AZ662093	AZ662093	1M0541C05
C	27	11	5.6	22	8	AZ324350	AZ324350	1M0046F14
C	28	11	5.6	24	9	CL653351	CL653351	PR10118b
C	29	11	5.6	24	9	CL653505	CL653505	PR10118b
C	30	11	5.6	25	8	AZ317218	AZ317218	1M0055J16
C	31	11	5.6	26	8	BZ767637	BZ767637	SALK_1391
C	32	11	5.6	27	1	AJ674038	AJ674038	AJ674038
C	33	11	5.6	28	1	AI699660	AI699660	weS8H01_x
C	34	11	5.6	28	6	CA797593	CA797593	Cac_BU_46
C	35	11	5.6	28	8	AQ025041	AQ025041	EP(2)1164
C	36	11	5.6	29	9	CG721624	CG721624	1119068C0
C	37	11	5.6	32	8	AZ360143	AZ360143	1M0103H18
C	38	11	5.6	33	1	AJ668521	AJ668521	AJ668521
C	39	11	5.6	33	7	R01622	R01622	ye79c10.61
C	40	11	5.6	33	9	DR732T	DR732T	DR732T
C	41	11	5.6	34	1	AI555208	AI555208	qT79a06_x
C	42	11	5.6	34	1	AI458585	AI458585	fmI4a03_x
C	43	11	5.6	34	8	AZ436025	AZ436025	1M0233C04
C	44	11	5.6	35	8	AZ491926	AZ491926	1M0025G18
C	45	11	5.6	35	9	BX532900	BX532900	AB16038

## ALIGNMENTS

RESULT 1	50 bp	linear	EST 28-JAN-2004
LOCUS	AU107206		
DEFINITION	AU107206 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone		
ACCESSION	HRC10935, mRNA sequence.		
VERSION	AU107206		
KEYWORDS	AU107206.1 GI:13556727		
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 50)		
AUTHORS	Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,		
	Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,		
	Sekaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.		
	Diverse transcriptional initiation revealed by fine, large-scale		
	mapping of mRNA start sites		
JOURNAL	EMBO Rep. 2 (5), 388-393 (2001)		
MEDLINE	21270072		
PUBMED	11375929		
COMMENT	Contact: Yutaka Suzuki		

## FEATURES

```

FEATURES
source
location/Qualifiers
1..50
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cd_name="HRC10935"
/clone_lib="Sugano Homo sapiens cDNA library"

```

ORIGIN

Query Match	7.1 %	Score	14	DB	1	Length	50
Best Local Similarity	100.0 %	Pred. No.	2.7e+04				
Matches	14	Conservative	0	Mismatches	0	Indels	0
						Gaps	0

QY	152	GGGACAGAGCAGA	165
Db	46	GGGACAGAGCAGA	33

RESULT 2  
LOCUS AU107207/c 50 bp mRNA linear EST 28-JAN-2004  
DEFINITION AU107207 Sugano Homo sapiens cDNA library Homo sapiens CDNA clone  
ACCESSION HRC12745, mRNA sequence.  
VERSION AU107207  
KEYWORDS AU107207.1 GI:13556728  
SOURCE EST. sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,  
Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,  
Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.  
TITLE Diverse transcriptional initiation revealed by fine, large-scale  
mapping of mRNA start sites  
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)  
MEDLINE 21270072  
PUBMED 11375929  
COMMENT Contact: Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: yusuzuki@ims.u-tokyo.ac.jp  
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and  
Sugano,S. Construction and characterization of a full  
length-enriched and a 5'-end-enriched CDNA library. Gene 200 (1-2),  
149-156 (1997).  
FEATURES  
source 1..50  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="HRC12745"  
/clone\_lib="Sugano Homo sapiens CDNA library"  
ORIGIN  
Query Match 7.1%; Score 14; DB 1; Length 50;  
Best Local Similarity 100.0%; Pred. No. 2.7e+04;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 152 GCGGACGAGCAGA 165  
DB 45 GCGGACGAGCAGA 32  
RESULT 3  
LOCUS AU107210/c 50 bp mRNA linear EST 28-JAN-2004  
DEFINITION AU107210 Sugano Homo sapiens cDNA library Homo sapiens CDNA clone  
ACCESSION KAT06101, mRNA sequence.  
VERSION AU107210  
KEYWORDS AU107210.1 GI:13556731  
SOURCE EST.  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,  
Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,  
Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.  
TITLE Diverse transcriptional initiation revealed by fine, large-scale  
mapping of mRNA start sites  
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)  
MEDLINE 21270072  
PUBMED 11375929  
COMMENT Contact: Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: yusuzuki@ims.u-tokyo.ac.jp  
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and  
Sugano,S. Construction and characterization of a full  
length-enriched and a 5'-end-enriched CDNA library. Gene 200 (1-2),  
149-156 (1997).  
FEATURES  
source 1..50  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="KAT06101"  
/clone\_lib="Sugano Homo sapiens CDNA library"  
ORIGIN  
Query Match 7.1%; Score 14; DB 1; Length 50;  
Best Local Similarity 100.0%; Pred. No. 2.7e+04;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 152 GCGGACGAGCAGA 165  
DB 49 GCGGACGAGCAGA 36  
RESULT 4  
LOCUS AU107211/c 50 bp mRNA linear EST 28-JAN-2004  
DEFINITION AU107211 Sugano Homo sapiens cDNA library Homo sapiens CDNA clone  
ACCESSION LMG05636, mRNA sequence.  
VERSION AU107211  
KEYWORDS AU107211.1 GI:13556732  
SOURCE EST.  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,  
Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,  
Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.  
TITLE Diverse transcriptional initiation revealed by fine, large-scale  
mapping of mRNA start sites  
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)  
MEDLINE 21270072  
PUBMED 11375929  
COMMENT Contact: Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: yusuzuki@ims.u-tokyo.ac.jp  
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and  
Sugano,S. Construction and characterization of a full  
length-enriched and a 5'-end-enriched CDNA library. Gene 200 (1-2),  
149-156 (1997).  
FEATURES  
source 1..50  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="LMG05636"  
/clone\_lib="Sugano Homo sapiens CDNA library"  
ORIGIN  
Query Match 7.1%; Score 14; DB 1; Length 50;  
Best Local Similarity 100.0%; Pred. No. 2.7e+04;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 152 GCGGACGAGCAGA 165  
DB 43 GCGGACGAGCAGA 30  
RESULT 5



```
US-10-719-900-56424/C
; Sequence 56424, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 56424
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-56424

Query Match
Best Local Similarity 100.0%; Pred. No. 1.5e+03; Length 25;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 CCCAGGAACTCAGA 104
DB 25 CCCAGGAACTCAGA 11

RESULT 3
US-10-719-900-195425/C
; Sequence 195425, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 195425
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-195425

Query Match
Best Local Similarity 100.0%; Pred. No. 1.5e+03; Length 25;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 GAACTCAGAGAGCG 109
DB 18 GAACTCAGAGAGCG 4

RESULT 4
US-10-719-900-489771/C
; Sequence 489771, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 489771
; LENGTH: 25
```

```
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-489771

Query Match
Best Local Similarity 100.0%; Pred. No. 1.5e+03; Length 25;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 AGCCCGAAGATGAG 41
DB 22 AGCCCGAAGATGAG 8

RESULT 5
US-10-719-900-955757
; Sequence 955757, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 955757
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-955757

Query Match
Best Local Similarity 100.0%; Pred. No. 1.5e+03; Length 25;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 CAGAGAGCGAGGCTG 115
DB 8 CAGAGAGCGAGGCTG 22

RESULT 6
US-10-719-956-180263/C
; Sequence 180263, Application US/10719956
; Publication No. US20040146910A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Rat
; FILE REFERENCE: 3527.1
; CURRENT APPLICATION NUMBER: US/10/719,956
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,836
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 699466
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 180263
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-10-719-956-180263

Query Match
Best Local Similarity 100.0%; Pred. No. 1.5e+03; Length 25;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 ACCTCAGAGAGCGAG 111
DB 23 ACCTCAGAGAGCGAG 9

RESULT 7
US-10-719-956-315460
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2005, 09:55:07 ; Search time 87.6346 Seconds  
(without alignments)  
3659.630 Million cell updates/sec

Title: US-09-887-941B-12

Perfect score: 196  
Sequence: 1 999CGCAGCAGGAGGAGTCA.....CTGAGACCGGAGGAGCAGCT 196

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1209694

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfillseq.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14	7.1	19	3	US-09-422-936-19
2	14	7.1	24	3	US-09-000-630C-17
3	14	7.1	24	3	US-08-862-730C-17
4	14	7.1	25	4	US-09-396-196G-23342
5	14	7.1	25	4	US-09-396-196G-112506
6	14	7.1	25	4	US-09-396-196G-112507
7	14	7.1	39	3	US-09-325-256-28
8	14	7.1	40	1	US-08-117-083-5
9	13	6.6	17	4	US-09-866-108A-2615
10	13	6.6	17	4	US-09-866-108A-2616
11	13	6.6	17	4	US-09-866-108A-2617
12	13	6.6	17	4	US-09-866-108A-2618
13	13	6.6	17	4	US-09-866-108A-2619
14	13	6.6	18	3	US-09-167-109-86
15	13	6.6	18	3	US-09-167-109-87
16	13	6.6	18	4	US-09-866-028-44
17	13	6.6	18	4	US-09-866-028-44
18	13	6.6	18	6	5512667-2
19	13	6.6	18	6	5512667-2
20	13	6.6	19	2	US-08-598-873-70
21	13	6.6	19	3	US-08-605-430-70
22	13	6.6	19	4	US-09-717-054-70
23	13	6.6	20	3	US-09-243-335-26
24	13	6.6	20	3	US-09-387-341-223
25	13	6.6	20	4	US-09-761-116-26
26	13	6.6	22	4	US-09-549-872B-17
27	13	6.6	25	4	US-09-866-108A-5544

28	13	6.6	25	4	US-09-866-108A-5545	Sequence 5545, Ap
29	13	6.6	25	4	US-09-866-108A-5546	Sequence 5546, Ap
30	13	6.6	25	4	US-09-866-108A-5547	Sequence 5547, Ap
31	13	6.6	25	4	US-09-866-108A-5548	Sequence 5548, Ap
32	13	6.6	25	4	US-09-866-108A-5549	Sequence 5549, Ap
33	13	6.6	25	4	US-09-866-108A-5550	Sequence 5550, Ap
34	13	6.6	25	4	US-09-866-108A-5551	Sequence 5551, Ap
35	13	6.6	25	4	US-09-866-108A-5552	Sequence 5552, Ap
36	13	6.6	25	4	US-09-866-108A-5553	Sequence 5553, Ap
37	13	6.6	25	4	US-09-866-108A-5554	Sequence 5554, Ap
38	13	6.6	25	4	US-09-866-108A-5555	Sequence 5555, Ap
39	13	6.6	25	4	US-09-866-108A-5556	Sequence 5556, Ap
40	13	6.6	25	4	US-09-396-196G-8224	Sequence 8224, Ap
41	13	6.6	25	4	US-09-396-196G-12058	Sequence 12058, A
42	13	6.6	25	4	US-09-396-196G-46116	Sequence 46116, A
43	13	6.6	25	4	US-09-396-196G-59077	Sequence 59077, A
44	13	6.6	25	4	US-09-396-196G-59088	Sequence 59088, A
45	13	6.6	25	4	US-09-396-196G-67675	Sequence 67675, A

ALIGNMENTS

RESULT 1  
US-09-422-936-19  
; Sequence 19, Application US/09422936  
; Patent No. 6465213  
; GENERAL INFORMATION:  
; APPLICANT: Ekstrand, Jonas  
; TITLE OF INVENTION: NEW NUCLEOTIDE SEQUENCES  
; FILE REFERENCE: 06275-165002  
; CURRENT APPLICATION NUMBER: US/09/422,936  
; PRIOR FILING DATE: 1999-10-22, 608  
; PRIOR APPLICATION NUMBER: US 09/242,608  
; PRIOR FILING DATE: 1999-02-19  
; PRIOR APPLICATION NUMBER: PCT/SE98/01947  
; PRIOR FILING DATE: 1998-10-27  
; PRIOR APPLICATION NUMBER: SWEDEN 9703914-2  
; PRIOR FILING DATE: 1997-10-27  
; PRIOR APPLICATION NUMBER: SWEDEN 9800864-2  
; PRIOR FILING DATE: 1998-03-16  
; PRIOR APPLICATION NUMBER: SWEDEN 9802575-2  
; PRIOR FILING DATE: 1998-07-17  
; NUMBER OF SEQ ID NOS: 85  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 19  
; LENGTH: 19  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
US-09-422-936-19

Query Match 7.1%; Score 14; DB 3; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 AGCCCGGAGAGG 131  
DB 6 AGCCCGGAGAGG 19

RESULT 2  
US-09-000-630C-17/c  
; Sequence 17, Application US/09000630C  
; Patent No. 6018029  
; GENERAL INFORMATION:  
; APPLICANT: Fuller, Gerald M  
; APPLICANT: Puentes, Nelson L.  
; TITLE OF INVENTION: DNA Encoding Canine Interleukin-1 Receptor  
; TITLE OF INVENTION: Antagonist  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Douglas C Murdock/ Bradley, Arant, Rose & White  
; STREET: 2001 Park Place, Suite 1400

CITY: Birmingham  
STATE: Alabama  
COUNTRY: USA  
ZIP: 35203-2736  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch,  
COMPUTER: IBM compatible  
OPERATING SYSTEM: Microsoft Windows  
SOFTWARE: Wordperfect 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/000,630C  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/862,730  
FILING DATE:  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 24 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: oligonucleotide  
US-09-000-630C-17

Query Match 7.1%; Score 14; DB 3; Length 24;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 109 GAGGCTGACAGGCC 122  
|||||  
DB 21 GAGGCTGACAGGCC 8

RESULT 3  
US-08-862-730C-17/c  
Sequence 17, Application US/08862730C  
Patent No. 6063600  
GENERAL INFORMATION:  
APPLICANT: Fuentes, Gerald M.  
TITLE OF INVENTION: DNA Encoding Canine Interleukin-1 Receptor  
TITLE OF INVENTION: Antagonist  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Douglas C Murdock/ Bradley, Arant, Rose & White  
STREET: 2001 Park Place, Suite 1400  
CITY: Birmingham  
STATE: Alabama  
COUNTRY: USA  
ZIP: 35203-2736  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch,  
COMPUTER: IBM compatible  
OPERATING SYSTEM: Microsoft Windows  
SOFTWARE: Wordperfect 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/862,730C  
FILING DATE: 5/23/97  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 24 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: oligonucleotide  
US-08-862-730C-17

Query Match 7.1%; Score 14; DB 3; Length 24;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 109 GAGGCTGACAGGCC 122  
|||||

DB 21 GAGGCTGACAGGCC 8

RESULT 4  
US-09-396-196G-23342/c  
Sequence 23342, Application US/09396196G  
Patent No. 6821724  
GENERAL INFORMATION:  
APPLICANT: Michael Miltmann  
APPLICANT: David Mack  
APPLICANT: David Lockhart  
APPLICANT: Affymetrix, Inc.  
TITLE OF INVENTION: Methods of Genetic Analysis  
FILE REFERENCE: 3101.1  
CURRENT APPLICATION NUMBER: US/09/396,196G  
CURRENT FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: 60/100,678  
PRIOR FILING DATE: 1998-09-17  
NUMBER OF SEQ ID NOS: 127806  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 23342  
LENGTH: 25  
TYPE: DNA  
ORGANISM: Mus musculus  
US-09-396-196G-23342

Query Match 7.1%; Score 14; DB 4; Length 25;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 110 AGGCTGACAGGCC 123  
|||||  
DB 16 AGGCTGACAGGCC 3

RESULT 5  
US-09-396-196G-112506/c  
Sequence 112506, Application US/09396196G  
Patent No. 6821724  
GENERAL INFORMATION:  
APPLICANT: Michael Miltmann  
APPLICANT: David Mack  
APPLICANT: David Lockhart  
APPLICANT: Affymetrix, Inc.  
TITLE OF INVENTION: Methods of Genetic Analysis  
FILE REFERENCE: 3101.1  
CURRENT APPLICATION NUMBER: US/09/396,196G  
CURRENT FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: 60/100,678  
PRIOR FILING DATE: 1998-09-17  
NUMBER OF SEQ ID NOS: 127806  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 112506  
LENGTH: 25  
TYPE: DNA  
ORGANISM: mus musculus  
US-09-396-196G-112506

Query Match 7.1%; Score 14; DB 4; Length 25;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 148 ACCAGCGACAGAG 161  
|||||  
DB 25 ACCAGCGACAGAG 12

RESULT 6  
US-09-396-196G-112507/c  
Sequence 112507, Application US/09396196G  
Patent No. 6821724  
GENERAL INFORMATION:  
APPLICANT: Michael Miltmann





CC of at least one target sequence. The method of analysis comprises  
 CC hybridizing at least one or more nucleic acids to at least two or more  
 CC nucleic acid probes and detecting the hybridisation. The nucleic acid  
 CC probes are attached to a solid support. The analysis comprises monitoring  
 CC gene expression levels, identifying allelic markers or polymorphisms,  
 CC or family members of a gene and a cross-species comparison. Each of the  
 CC nucleic acids further comprises a tag sequence. The array of nucleic acid  
 CC probes is useful in situ hybridisation, in Southern, Northern or dot-  
 CC blot hybridisation to identify or detect the sequence or specific  
 CC mutations of any gene, in mapping the 5' termini of mRNA molecules by  
 CC primer extensions or in screening cDNA or genomic libraries or subclones  
 CC for additional subclones containing segments of DNA that have been  
 CC isolated and previously sequenced. The sequence presented is one of the  
 CC nucleic acid probes incorporated in the microarray. Note: The sequence  
 CC data for this patent can also be obtained in electronic format directly  
 CC from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html)

XX Sequence 25 BP; 8 A; 6 C; 9 G; 2 T; 0 U; 0 Other;

Query Match 7.7%; Score 15; DB 9; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 92 CAGGAACCTCAGAGA 106  
 |||||  
 DB 3 CAGGAACCTCAGAGA 17

RESULT 2  
 ABA01159  
 ID ABA01159 standard; DNA; 33 BP.

XX ABA01159;

XX 24-JAN-2002 (first entry)

DE Human zinc finger protein 72 PCR primer #3.

XX Human; zinc finger protein 72; cytosolic; virucidal; immunomodulatory;

KW antiinflammatory; haemostatic; anti-HIV; gene therapy; cancer;

KM haemopathy; HIV infection; immunological disease; inflammation;

KW nervous system disease; developmental disorder; PCR primer; se.

XX Homo sapiens.

XX WO200174866-A1.

XX 11-OCT-2001.

XX 26-FEB-2001; 2001WO-CN000196.

XX 02-MAR-2000; 2000CN-00111866.

XX (BIOW-) B10W10W GENE DEV INC SHANGHAI.

XX Mao Y, Xie Y;

XX WPI; 2002-025795/03.

XX New human zinc finger protein 72 for diagnosing and treating malignant  
 XX tumor, hemopathy, human immunodeficiency virus infection, immunological  
 XX diseases and various inflammations.

XX Example 5; Page 13; 38pp; Chinese.

XX The present invention relates to human zinc finger protein 72 (see  
 CC AAM5332). The zinc finger protein and its coding sequence are useful in  
 CC the diagnosis and treatment of cancer, haemopathy, HIV infection,  
 CC immunological diseases, various inflammations, nervous system diseases  
 CC and developmental disorders. The present sequence is a PCR primer, which  
 CC was used in an example from the present invention  
 XX Sequence 33 BP; 9 A; 8 C; 9 G; 7 T; 0 U; 0 Other;

Query Match 7.7%; Score 15; DB 6; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 85 TGAGTCCAGGAACC 99  
 |||||  
 DB 15 TGAGTCCAGGAACC 29

RESULT 3  
 AAQ10273  
 ID AAQ10273 standard; DNA; 38 BP.

XX AAQ10273;

XX 25-MAR-2003 (revised)

XX 03-APR-1991 (first entry)

DE Probe B(II) for bovine BMP exon 3.

XX Bone morphogenetic protein; cartilage; antibodies; osteoporosis;

KW osteoarthritis; screening.

XX Synthetic.

XX EP409472-A.

XX 23-JAN-1991.

XX 11-JUL-1990; 90EP-00307568.

XX 19-JUL-1989; 89US-00382805.

XX (CHIR ) CHIRON CORP.

XX Keifer MC, Maslarsz FR, Barr PJ;

XX WPI; 1991-024045/04.

XX Purified mammalian bone morphogenetic protein - and recombinant DNA  
 XX encoding it, useful in pharmaceuticals for cartilage and bone formation.

XX Disclosure; Fig 1; 26pp; English.

XX The sequence is the complement of probe B designed to bind to a DNA  
 XX encoding tryptic peptide B, found in exon 3 of bovine BMP genomic DNA.

XX Recombinantly produced BMP can be used to induce cartilage and bone  
 XX formation in vertebrates (dose =ca. 0.1 ug-100mg/kg). Ads raised to the

XX protein may be used to inhibit or to reverse e.g. osteoporosis,  
 XX osteoarthritis, etc. Purified BMP will be partic. useful in the design

XX and screening of cartilage/bone growth inhib- itors. See also AAQ10265-  
 XX Q10276. (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 38 BP; 7 A; 11 C; 11 G; 5 T; 0 U; 4 Other;

Query Match 7.7%; Score 15; DB 2; Length 38;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 103 GAGAGCGAGGCTGAC 117  
 |||||  
 DB 1 GAGAGCGAGGCTGAC 15

RESULT 4  
 ABA01162  
 ID ABA01162 standard; DNA; 41 BP.

XX ABA01162;

XX 24-JAN-2002 (first entry)



sequence differences using ligase detection reaction  
 Patent: WO 0179548-A 8939 25-OCT-2001;  
 CORNELL RESEARCH FOUNDATION, INC. (US)

JOURNAL  
 FEATURES  
 source  
 1..20  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="Hypothetical Probe Sequence"

ORIGIN  
 Query Match 7.1%; Score 14; DB 6; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+04;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 26 GAGCCCGAAGATGG 39  
 DB 4 GAGCCCGAAGATGG 17  
 |||||  
 |||||

RESULT 3  
 E13893 24 bp DNA linear PAT 27-APR-1998  
 LOCUS PCR primer for gaining human G protein-coupling receptor gene.  
 DEFINITION E13893  
 ACCESSION E13893  
 VERSION JP 1997238686-A/2.  
 KEYWORDS  
 SOURCE unidentified  
 ORGANISM unidentified.

REFERENCE  
 1 (bases 1 to 24)  
 Hinuma, K. and Fujii, A.  
 NEW G-PROTEIN COUPLED TYPE RECEPTOR PROTEIN, ITS PRODUCTION AND  
 TITLE Patent: JP 1997238686-A 2 16-SEP-1997;  
 JOURNAL TAKEDA CHEM IND LTD

COMMENT  
 OS None  
 OC Artificial sequences.  
 PN JP 1997238686-A/2  
 PD 16-SEP-1997  
 PF 07-MAR-1996 JP 1996050678  
 PI HINUMA KUNIJU, FUJII AKIRA  
 PC C12N15/09,A61K48/00,C07H21/04,C12N1/21,C12P21/02,C12P21/08, PC  
 C12Q1/02,  
 PC G01N33/566//A61K39/395,(C12N1/21,C12R1.19),(C12P21/02, PC  
 C12R1.19),(C12P21/08,  
 PC C12R1.91);  
 CC strandedness: Single;  
 CC topology: Linear;  
 CC hypothetical: No;  
 CC anti-sense: No;  
 FH key Location/Qualifiers  
 FH source 1..24  
 FT Location/Qualifiers  
 FT source 1..24  
 1..24  
 /organism="Artificial sequences".  
 /organism="unidentified"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32644"

ORIGIN  
 Query Match 7.1%; Score 14; DB 6; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 3e+04;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 83 GATGAGTCCCGAGA 96  
 DB 17 GATGAGTCCCGAGA 4  
 |||||  
 |||||

RESULT 4  
 AX292544 24 bp DNA linear PAT 21-NOV-2001  
 LOCUS AX292544

DEFINITION Sequence 4306 from Patent WO0179548.  
 ACCESSION AX292544  
 VERSION AX292544.1 GI:17054227  
 KEYWORDS  
 SOURCE synthetic construct  
 ORGANISM synthetic construct  
 other sequences; artificial sequences.

REFERENCE  
 1  
 Barany, F., Zilvi, M., Gerry, N.P., Favis, R. and Kliman, R.  
 TITLE Method of designing addressable array for detection of nucleic acid  
 JOURNAL Sequence differences using ligase detection reaction  
 Patent: WO 0179548-A 4306 25-OCT-2001;  
 CORNELL RESEARCH FOUNDATION, INC. (US)  
 FEATURES  
 source  
 1..24  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="Hypothetical Probe Sequence"

ORIGIN  
 Query Match 7.1%; Score 14; DB 6; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 3e+04;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 26 GAGCCCGAAGATGG 39  
 DB 4 GAGCCCGAAGATGG 17  
 |||||  
 |||||

RESULT 5  
 AX318212 24 bp DNA linear PAT 14-DEC-2001  
 LOCUS Sequence 13 from Patent WO0190163.  
 DEFINITION AX318212  
 ACCESSION AX318212  
 VERSION AX318212.1 GI:17900895  
 KEYWORDS  
 SOURCE synthetic construct  
 ORGANISM synthetic construct  
 other sequences; artificial sequences.

REFERENCE  
 1  
 Ng, G.Y.  
 TITLE Use of the gaba b<sub>2</sub> receptor in assays to identify gamma  
 agonists  
 JOURNAL hydroxybutyrate agonists, antagonists, and allosteric modulators of  
 Patent: WO 0190163-A 13 29-NOV-2001;  
 Merck Froese Canada & Co. (CA)  
 FEATURES  
 source  
 1..24  
 Location/Qualifiers  
 1..24  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="PCR primer"

ORIGIN  
 Query Match 7.1%; Score 14; DB 6; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 3e+04;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 118 AGCCCGGAGAG 131  
 DB 3 AGCCCGGAGAG 16  
 |||||  
 |||||

RESULT 6  
 AX815614 25 bp DNA linear PAT 09-DEC-2003  
 LOCUS AX815614  
 DEFINITION Sequence 12 from Patent WO03066904.  
 ACCESSION AX815614  
 VERSION AX815614.1 GI:39646311  
 KEYWORDS  
 SOURCE synthetic construct  
 ORGANISM synthetic construct

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2005, 15:23:04 ; Search time 1905.96 Seconds  
(without alignments)  
4393.656 Million cell updates/sec

Title: US-09-887-941B-11

Perfect score: 220  
Sequence: 1 ctgctctctctctctctt.....ccactgcgcgcgcgcgcgc 220

Scoring table:  
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 159776

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16	7.3	40	9	CL437571 PST5874-N
2	14	6.4	23	1	AB094450 AB094450
3	14	6.4	28	9	CG722869 1119073G0
4	14	6.4	29	8	BH755587 SALK_0518
5	14	6.4	31	1	AJ732952 AJ732952
6	14	6.4	45	9	CL658678 PR10132a
7	14	6.4	47	7	CF310621 ABE--05-F
8	14	6.4	47	8	AZ513870 1M0360X08
9	14	6.4	49	9	CU529376 HTY40G02.
10	14	6.4	50	1	AU102598 AU102598
11	14	6.4	50	1	AU107543 AU107543
12	14	6.4	50	9	BX288916 Arabidops
13	14	6.4	50	9	CG710737 1119018F0
14	13	5.9	19	7	CF324209 HDN--05-O
15	13	5.9	26	8	AZ454352 1M0256G18
16	13	5.9	27	8	AZ582689 1M0376D13
17	13	5.9	29	1	AU256240 AU256240
18	13	5.9	29	7	N22525
19	13	5.9	29	8	AZ356943 1M0098D16
20	13	5.9	32	8	AZ828692 2M0105N15
21	13	5.9	32	8	AZ408275 1M0179G20
22	13	5.9	32	7	BX656232 Arabidops
23	13	5.9	33	7	CF307411 HDX1--06-
24	13	5.9	36	5	BQ035387 SK2-0117

25.	13	5.9	43	8	AZ830331	AZ830331 2M0109H03
26	13	5.9	43	8	AZ877472	AZ877472 2M0192L19
27	13	5.9	46	1	A1631069	A1631069 t232b10.x
28	13	5.9	46	8	CC178512	CC178512 NPX329_Ba
29	13	5.9	47	8	AZ420561	AZ420561 1M0198021
30	13	5.9	47	9	AJ591522	AJ591522 Arabidops
31	13	5.9	48	8	BH901106	BH901106 KG09864-5
32	13	5.9	49	8	AZ642798	AZ642798 1M0506D02
33	13	5.9	49	9	CR356331	CR356331 Arabidops
34	13	5.9	50	1	AU104216	AU104216 AU104216
35	13	5.9	50	1	AU104712	AU104712 AU104712
36	13	5.9	50	1	AU107973	AU107973 AU107973
37	13	5.9	50	9	AJ591492	AJ591492 Arabidops
38	13	5.9	50	9	AL943250	AL943250 Arabidops
39	12	5.5	20	8	AZ490328	AZ490328 1M0323A18
40	12	5.5	25	1	A1539240	A1539240 tp64b08.x
41	12	5.5	25	8	AZ635993	AZ635993 1M0493520
42	12	5.5	25	8	BZ762357	BZ762357 SALK_1003
43	12	5.5	31	8	AZ336296	AZ336296 1M0066F01
44	12	5.5	31	8	AZ763536	AZ763536 1M0559I02
45	12	5.5	31	9	CG707251	CG707251 1119001G0

#### ALIGNMENTS

RESULT 1  
CL437571 40 bp DNA linear GSS 18-MAR-2004  
PST5874-NL.Seg MICB1 Mus musculus genomic clone PST5874-NL.Seg.  
DEFINITION  
CL437571  
ACCESSION  
CL437571 GI:45573360  
VERSION  
CL437571.1  
KEYWORDS  
GSS.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 (bases 1 to 40)  
Hicks, G.G.  
AUTHORS  
www.Bscellis.ca  
TITLE  
Unpublished (2002)  
JOURNAL  
Contact: Hicks GG  
COMMENT  
Mammalian Functional Genomics Centre  
Manitoba Institute of Cell Biology, University of Manitoba  
ON5029, 675 McDermot Ave, Winnipeg, MB R3E 0V9, Canada  
Tel: 204 787 2133  
Fax: 204 787 2190  
Email: hicksg@cc.umanitoba.ca

U3NeosVI gene trap. Tag generated by plasmid rescue. Additional  
sequence information and target gene cloning can be generated. ES  
cell line harboring insertion mutation of target gene is available.  
Sequence analysis available from  
http://140.193.242.7/esdb/public\_search\_frame.php?PST=PST5874-NL.Se  
q  
Class: Gene Trap.  
Location/Qualifiers  
1..40  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="129 sv"  
/db\_xref="taxon:10090"  
/clone="PST5874-NL.Seg"  
/sex="Male"  
/cell\_type="Embryonic stem cell"  
/cell\_line="D3H (J1 subclone)"  
/clone\_1b="MICB1"  
/note="Vector: U3NeosVI"

#### FEATURES

source  
ORIGIN  
Query Match 7.3%; Score 16; DB 9; Length 40;  
Best Local Similarity 100.0%; Pred. No. 2.5e+03;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 167 CCCGAATCTGCAGGT 182  
 |||||  
 Db 8 CCCGAATCTGCAGGT 23

## RESULT 2

AB094450 23 bp mRNA linear EST 01-MAY-2003  
 LOCUS AB094450.1 lambda Triplex2 rice phloem sap cDNA Oryza sativa  
 DEFINITION (japonica cultivar-group) cDNA clone PA158, mRNA sequence.  
 AB094450  
 VERSION AB094450.1 GI:30307355  
 KEYWORDS EST.

ACCESION Oryza sativa (japonica cultivar-group)  
 ORIGIN Oryza sativa (japonica cultivar-group)

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.  
 1 (bases 1 to 23)

AUTHORS Mano, H., Noguchi, M., Oshima, T., Yoneyama, T., Hayashi, H. and  
 Fujikawa, T.  
 TITLE Small RNAs detected in the rice phloem sap  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Hiromori Mano  
 Plant Genome Center Co., Ltd  
 Nara 630-0192, Tsukuba, Ibaraki 305-0856, Japan  
 Tel: 81-298-39-4823  
 Email: hmano@pccdn.co.jp.

FEATURES  
 source Location/Qualifiers

1..23

/organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="PA158"  
 /tissue\_type="phloem"  
 /clone\_id="lambda Triplex2 rice phloem sap cDNA"

## ORIGIN

Query Match 6.4%; Score 14; DB 1; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+04;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 TTTCCTCTCTCTC 38  
 |||||  
 Db 20 TTTCCTCTCTCTC 7

RESULT 3  
 CG722869 28 bp DNA linear GSS 20-OCT-2003  
 LOCUS 1119073G09.1EB\_x1 1119 - Rescuedu Grid AA Zea mays genomic, genomic  
 DEFINITION survey sequence.  
 CG722869  
 VERSION CG722869.1 GI:37758149  
 KEYWORDS GSS.

ACCESION Zea mays  
 ORIGIN Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoidae; Andropogoneae; Zea.  
 1 (bases 1 to 28)

AUTHORS Walbot, V.  
 TITLE Maize genomic sequences found using engineered Rescuedu transposon  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Walbot V  
 Department of Biological Sciences  
 Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 725 8221  
 Email: walbot@stanford.edu

Very probable ligation site of ends cut by single endonuclease.  
 Reverse complemented post-ligation sequence from source sequence.  
 Plate: 1119073 row: G column: 09  
 Class: transposon-tagged.  
 Location/Qualifiers

FEATURES  
 source

1..28

/organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /cultivar="mixed background W23/A18/B73/K55"  
 /db\_xref="taxon:4577"  
 /tissue\_type="leaf"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"

/clone\_id="1119 - Rescuedu Grid AA"  
 /note="Organ: leaf; Vector: Rescuedu (engineered from  
 pBluescript backbone); Site 1: BamHI; Site 2: BglII;  
 Rescuedu is a 4.9 kb, modified maize Mu transposon  
 designed to allow plasmid rescue from total genomic DNA.  
 Mu elements insert preferentially into transcription  
 units. For more information on Rescuedu, go to the web  
 site 'www.zmdb.iastate.edu' and follow the links for  
 'Rescuedu', 'Grid AA' was grown at UC San Diego in 2002. DNA  
 was extracted from leaf strips, double digested using  
 BamHI and BglII, and ligated to form circular plasmids.  
 DH10B cells were transformed and then screened on LB  
 plates with ampicillin."

## ORIGIN

Query Match 6.4%; Score 14; DB 9; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+04;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 CTTTCCCTCTCC 29  
 |||||  
 Db 3 CTTTCCCTCTCC 16

RESULT 4  
 BH755587 29 bp DNA linear GSS 01-MAR-2002  
 LOCUS SALX\_051892.52.10 x Arabidopsis thaliana TDNA insertion lines  
 DEFINITION Arabidopsis thaliana genomic clone SALX\_051892.52.10.x, genomic  
 survey sequence.  
 BH755587  
 VERSION BH755587.1 GI:19035884  
 KEYWORDS GSS.

ACCESION Arabidopsis thaliana (thale cress)  
 ORIGIN Arabidopsis thaliana

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 29)

AUTHORS Alonso, J.M., Leisner, T.J., Barajas, P., Chen, H., Cheuk, R.,  
 Gadrinab, C., Jeske, A., Karnes, W., Kim, C.J., Parker, H., Prednis, L.,  
 Shinn, P., Zimmerman, J., and Ecker, J.R.  
 TITLE A sequence-indexed library of insertion mutations in the  
 Arabidopsis Genome  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Joseph R. Ecker  
 Salk Institute Genomic Analysis Laboratory (SIGNAL)  
 The Salk Institute for Biological Studies  
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
 Tel: 858 453 4100 x1752  
 Fax: 858 558 6379  
 Email: ecker@salk.edu

This is single pass sequence recovered from the left border of  
 TDNA. This sequence lies within 300 bases of the 5' end of  
 At2g38080.  
 Class: TDNA tagged.  
 Location/Qualifiers

FEATURES  
 source

1..29

/organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2005, 12:40:34 ; Search time 452.163 Seconds  
(without alignments)  
3410.666 Million cell updates/sec

Title: US-09-887-941B-11

Perfect score: 220  
Sequence: 1 ctgctctctctctctctt.....ccactcgccgccccgcgcgc 220

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 8765912 seqs, 3504951483 residues

Word size : 0

Total number of hits satisfying chosen parameters: 11015406

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Listing first 45 summaries

Database :

Published Applications NA:  
1: /cgn2\_6/prodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/prodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/prodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/prodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/prodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/prodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/prodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/prodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/prodata/2/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/prodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/prodata/2/pubpna/US09\_PUBCOMB.seq:\*  
12: /cgn2\_6/prodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:\*  
14: /cgn2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:\*  
15: /cgn2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:\*  
16: /cgn2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:\*  
17: /cgn2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:\*  
18: /cgn2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:\*  
19: /cgn2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:\*  
20: /cgn2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:\*  
21: /cgn2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:\*  
22: /cgn2\_6/prodata/2/pubpna/US11\_PUBCOMB.seq:\*  
23: /cgn2\_6/prodata/2/pubpna/US11\_PUBCOMB.seq:\*  
24: /cgn2\_6/prodata/2/pubpna/US11\_PUBCOMB.seq:\*  
25: /cgn2\_6/prodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
26: /cgn2\_6/prodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	25	11.4	25	US-10-719-900-365536	Sequence 365536, A
2	25	11.4	25	US-10-719-900-616433	Sequence 616433, A
3	25	11.4	25	US-10-719-900-654770	Sequence 654770, A
4	25	11.4	25	US-10-719-900-717050	Sequence 717050, A
5	25	11.4	25	US-10-719-900-958924	Sequence 958924, A
6	25	11.4	25	US-11-036-317-20600	Sequence 20600, A
7	25	11.4	25	US-11-036-317-23369	Sequence 23369, A

8	25	11.4	25	US-11-036-317-45840	Sequence 45840, A
9	25	11.4	25	US-11-036-317-70907	Sequence 70907, A
10	25	11.4	25	US-11-036-317-121383	Sequence 121383, A
11	25	11.4	25	US-11-036-317-595646	Sequence 595646, A
12	25	11.4	25	US-11-036-317-610660	Sequence 610660, A
13	25	11.4	25	US-11-036-317-721050	Sequence 721050, A
14	25	11.4	25	US-11-036-317-846045	Sequence 846045, A
15	24	10.9	24	US-10-479-306-47	Sequence 47, App1
16	21	9.5	25	US-10-719-900-508538	Sequence 508538, A
17	21	9.5	25	US-11-036-317-34675	Sequence 34675, A
18	21	9.5	25	US-11-036-317-54334	Sequence 54334, A
19	21	9.5	25	US-11-036-317-665140	Sequence 665140, A
20	18	8.2	25	US-11-036-317-332917	Sequence 332917, A
21	18	8.2	25	US-11-036-317-953540	Sequence 953540, A
22	17	7.7	25	US-10-719-900-810729	Sequence 810729, A
23	17	7.7	25	US-11-036-317-55915	Sequence 55915, A
24	17	7.7	25	US-11-036-317-135978	Sequence 135978, A
25	17	7.7	25	US-11-036-317-261947	Sequence 261947, A
26	17	7.7	25	US-11-036-317-360730	Sequence 360730, A
27	17	7.7	25	US-11-036-317-771422	Sequence 771422, A
28	17	7.7	25	US-11-036-317-969552	Sequence 969552, A
29	16	7.3	25	US-10-719-900-310911	Sequence 310911, A
30	16	7.3	25	US-10-719-900-699479	Sequence 699479, A
31	16	7.3	25	US-10-956-157-108671	Sequence 108671, A
32	16	7.3	25	US-10-956-157-108672	Sequence 108672, A
33	16	7.3	25	US-10-956-157-108674	Sequence 108674, A
34	16	7.3	25	US-10-956-157-108675	Sequence 108675, A
35	16	7.3	25	US-10-956-157-108681	Sequence 108681, A
36	16	7.3	25	US-10-956-157-108682	Sequence 108682, A
37	16	7.3	25	US-11-036-317-3223	Sequence 3223, App
38	16	7.3	25	US-11-036-317-13946	Sequence 13946, A
39	16	7.3	25	US-11-036-317-181263	Sequence 181263, A
40	16	7.3	25	US-11-036-317-244776	Sequence 244776, A
41	16	7.3	25	US-11-036-317-510859	Sequence 510859, A
42	16	7.3	25	US-11-036-317-563387	Sequence 563387, A
43	16	7.3	25	US-11-036-317-866279	Sequence 866279, A
44	16	7.3	30	US-10-813-977-35	Sequence 35, App1
45	15	6.8	20	US-10-188-404-39	Sequence 39, App1

#### ALIGNMENTS

RESULT 1  
US-10-719-900-365536  
; Sequence 365536, Application US/10719900  
; Publication No. US20050026164A1  
; GENERAL INFORMATION:  
; APPLICANT: Xue Mei Zhou  
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse  
; FILE REFERENCE: 3528.1  
; CURRENT APPLICATION NUMBER: US/10/719, 900  
; PRIOR FILING DATE: 2003-11-20  
; PRIOR APPLICATION NUMBER: 60/427, 808  
; NUMBER OF SEQ ID NOS: 982914  
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1  
; SEQ ID NO 365536  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-719-900-365536

Query Match 11.4%; Score 25; DB 21; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 CGTGGACATTTCGGAATTTCGA 179  
DB 1 CGTGGACATTTCGGAATTTCGA 25

RESULT 2

```
US-10-719-900-616433
; Sequence 616433, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; PRIOR FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 616433
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-616433

Query Match      11.4%; Score 25; DB 21; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      152 GCGCGTGGACATTTCCCGAATTCT 176
|||
Db      1 GCGCGTGGACATTTCCCGAATTCT 25

RESULT 3
US-10-719-900-654770
; Sequence 654770, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; PRIOR FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 654770
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-654770

Query Match      11.4%; Score 25; DB 21; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      76 GGTGCTTTGTCAACGGCGGCCAC 100
|||
Db      1 GGTGCTTTGTCAACGGCGGCCAC 25

RESULT 4
US-10-719-900-717050
; Sequence 717050, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; PRIOR FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 717050
; LENGTH: 25
```

```
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-717050

Query Match      11.4%; Score 25; DB 21; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      77 GTGCTTTGTCAACGGCGGCCACT 101
|||
Db      1 GTGCTTTGTCAACGGCGGCCACT 25

RESULT 5
US-10-719-900-958924
; Sequence 958924, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; PRIOR FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 958924
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-958924

Query Match      11.4%; Score 25; DB 21; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      83 TTGTCAACGGCGGCACCTGCCGA 107
|||
Db      1 TTGTCAACGGCGGCACCTGCCGA 25

RESULT 6
US-11-036-317-20600
; Sequence 20600, Application US/11036317
; Publication No. US20050214823A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Alan
; TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
; FILE REFERENCE: 3654.1
; CURRENT APPLICATION NUMBER: US/11/036,317
; PRIOR FILING DATE: 2005-01-13
; PRIOR APPLICATION NUMBER: US 60/536,639
; NUMBER OF SEQ ID NOS: 991174
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 20600
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-036-317-20600

Query Match      11.4%; Score 25; DB 24; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      150 CCGCGGTGGACATTTCCCGAATT 174
|||
Db      1 CCGCGGTGGACATTTCCCGAATT 25

RESULT 7
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2005, 09:55:07 ; Search time 98.1654 Seconds  
(without alignments)  
3659.630 Million cell updates/sec

Title: US-09-887-941B-11

Perfect score: 220

Sequence: 1 ctgcctctctctctctctctt.....ccactcgccgccccgcgcgc 220

Scoring table: OLIGO NUC

Gapop\_60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1209694

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Listing first 45 summaries

Database :

Issued\_Patents\_NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysts of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	9.1	38	2	US-08-857-946-149
2	20	9.1	38	3	US-08-970-740-149
3	20	9.1	40	2	US-08-857-946-150
4	20	9.1	40	3	US-08-970-740-150
5	16	7.3	34	1	US-08-413-813-41
6	16	7.3	34	2	US-08-467-146-41
7	15	6.8	20	3	US-08-275-951-39
8	15	6.8	31	3	US-08-848-373-1
9	15	6.8	36	1	US-08-370-567-30
10	15	6.8	36	1	US-08-438-759-30
11	15	6.8	36	1	US-08-538-911-5
12	15	6.8	36	1	US-08-117-361C-17
13	15	6.8	36	5	PCT-US94-05591-5
14	15	6.8	36	5	PCT-US94-05684-30
15	14	6.4	20	3	US-09-357-073-12
16	14	6.4	20	3	US-09-357-073-13
17	14	6.4	20	3	US-09-357-073-14
18	14	6.4	20	3	US-09-357-073-15
19	14	6.4	25	4	US-09-396-196G-2778
20	14	6.4	25	4	US-09-396-196G-25595
21	14	6.4	25	4	US-09-396-196G-66137
22	14	6.4	25	4	US-09-396-196G-66137
23	14	6.4	27	3	US-09-253-396A-59
24	14	6.4	29	1	US-08-342-411A-32
25	14	6.4	29	5	PCT-US94-12883-32
26	14	6.4	30	3	US-07-861-458C-61
27	14	6.4	33	1	US-08-036-555B-90

28	14	6.4	33	1	US-08-036-555B-93	Sequence 93, Appl
29	14	6.4	33	1	US-08-469-569-90	Sequence 90, Appl
30	14	6.4	33	1	US-08-469-569-93	Sequence 93, Appl
31	14	6.4	33	1	US-08-249-322A-90	Sequence 90, Appl
32	14	6.4	33	1	US-08-249-322A-93	Sequence 93, Appl
33	14	6.4	33	1	US-08-469-526A-90	Sequence 90, Appl
34	14	6.4	33	1	US-08-469-526A-93	Sequence 93, Appl
35	14	6.4	33	2	US-08-734-591A-90	Sequence 90, Appl
36	14	6.4	33	2	US-08-734-591A-93	Sequence 93, Appl
37	14	6.4	33	2	US-08-469-660-90	Sequence 90, Appl
38	14	6.4	33	2	US-08-469-660-93	Sequence 93, Appl
39	14	6.4	33	2	US-08-671-978A-16	Sequence 16, Appl
40	14	6.4	33	3	US-08-470-335-90	Sequence 90, Appl
41	14	6.4	33	3	US-08-470-335-93	Sequence 93, Appl
42	14	6.4	33	3	US-08-735-021-90	Sequence 90, Appl
43	14	6.4	33	3	US-08-735-021-93	Sequence 93, Appl
44	14	6.4	33	3	US-08-734-664A-90	Sequence 90, Appl
45	14	6.4	33	3	US-08-734-664A-93	Sequence 93, Appl

#### ALIGNMENTS

RESULT 1  
US-08-857-946-149  
Sequence 149, Application US/08857946  
Patent No. 5994075  
GENERAL INFORMATION:  
APPLICANT: Goodfellow, P.N.  
TITLE OF INVENTION: METHODS FOR IDENTIFYING A MUTATION IN A  
TITLE OF INVENTION: GENE OF INTEREST  
NUMBER OF SEQUENCES: 162  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: Banner & Witcoff, Inc.  
STREET: 75 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1807  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/857,946  
FILING DATE: 16-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/60/017,824  
FILING DATE: 17-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Kathleen M. Williams  
REGISTRATION NUMBER: 34,380  
REFERENCE/DOCKET NUMBER: 3529/05573  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-345-9100  
TELEFAX: 617-345-9111  
INFORMATION FOR SEQ ID NO: 149:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 38 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
FEATURE:  
NAME/KEY: primer Pax6mm121f  
US-08-857-946-149

Query Match 9.1%, Score 20; DB 2; Length 38;  
Best Local Similarity 100.0%; Pred. No. 2.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



OY 50 GTCACAGCGAGTGAATCAG 69  
|||||  
DB 19 GTCACAGCGAGTGAATCAG 38

## RESULT 2

US-08-970-740-149  
; Sequence 149, Application US/08970740  
; Patent No. 6015670  
; GENERAL INFORMATION:  
; APPLICANT: Goodfellow, P.N.  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING A MUTATION IN A  
; TITLE OF INVENTION: GENE OF INTEREST  
; NUMBER OF SEQUENCES: 162  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Witcoff, Inc.  
; STREET: 28 State Street, 28th Floor  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Wordperfect 6.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/970,740  
; FILING DATE: 14-NOV-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/857,946  
; FILING DATE: 16-MAY-1997  
; PRIOR APPLICATION NUMBER:  
; APPLICATION NUMBER: 60/017,824  
; FILING DATE: 17-MAY-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kathleen M. Williams  
; REGISTRATION NUMBER: 34,380  
; REFERENCE/DOCKET NUMBER: 3529/59829  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-227-4399  
; INFORMATION FOR SEQ ID NO: 149:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 38 bases  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: other nucleic acid  
; FEATURE:  
; NAME/KEY: primer Pax6mm121f  
US-08-970-740-149

Query Match 9.1%; Score 20; DB 3; Length 38;  
Best Local Similarity 100.0%; Pred. No. 2.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 50 GTCACAGCGAGTGAATCAG 69  
|||||  
DB 19 GTCACAGCGAGTGAATCAG 38

RESULT 3  
US-08-857-946-150/c  
; Sequence 150, Application US/08857946  
; Patent No. 5994075  
; GENERAL INFORMATION:  
; APPLICANT: Goodfellow, P.N.  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING A MUTATION IN A  
; TITLE OF INVENTION: GENE OF INTEREST  
; NUMBER OF SEQUENCES: 162  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Witcoff, Inc.

STREET: 75 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA

ZIP: 02109-1807

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Wordperfect 6.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/857,946

FILING DATE: 16-MAY-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/60/017,824

FILING DATE: 17-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Kathleen M. Williams

REGISTRATION NUMBER: 34,380

REFERENCE/DOCKET NUMBER: 3529/05573

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-345-9100

INFORMATION FOR SEQ ID NO: 150:

SEQUENCE CHARACTERISTICS:

LENGTH: 40 bases

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid

FEATURE:

NAME/KEY: primer Pax6mm121r

US-08-857-946-150

Query Match 9.1%; Score 20; DB 2; Length 40;  
Best Local Similarity 100.0%; Pred. No. 2.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 160 GACATTTCCCGAATTCGCA 179  
|||||  
DB 40 GACATTTCCCGAATTCGCA 21

## RESULT 4

US-08-970-740-150/c  
; Sequence 150, Application US/08970740  
; Patent No. 6015670  
; GENERAL INFORMATION:  
; APPLICANT: Goodfellow, P.N.  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING A MUTATION IN A  
; TITLE OF INVENTION: GENE OF INTEREST  
; NUMBER OF SEQUENCES: 162  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Witcoff, Inc.  
; STREET: 28 State Street, 28th Floor  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Wordperfect 6.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/970,740  
; FILING DATE: 14-NOV-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/857,946  
; FILING DATE: 16-MAY-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/017,824



CC to tissue-disruption to provide a mixed population (MP) comprising UC,  
 CC subjecting MP to a cell size-discrimination (SD) step, and simultaneously  
 CC or sequentially with SD, subjecting the cell population obtained to a  
 CC cell-surface marker-discrimination step. Also described: (1) a  
 CC substantially homogeneous population of undifferentiated cells (I)  
 CC prepared by (M); (2) a composition (II) for use in cell replacement  
 CC therapy, comprising a population of substantially homogeneous population  
 CC of neural stem cells (NSCs) generated by (M); and (3) a composition (III)  
 CC comprising a growth factor identified using a homogeneous population of  
 CC NSCs generated by (M). (I) can have neurotropic, neuroprotective and  
 CC antiapoptotic activities, and can be used in gene therapy. (M) is  
 CC useful for generating a substantially homogeneous population of  
 CC undifferentiated cells such as NSCs from a biological sample, and is  
 CC useful for the replacement of neural or non-neural tissue in an animal.  
 CC (II) is useful in cell replacement therapy in an organ such as the brain  
 CC or in the nervous system, preferably central nervous system (CNS), for  
 CC treating a CNS disorder such as Alzheimer's disease, Parkinson's disease,  
 CC acute brain injury and CNS dysfunction. (I) is useful for the repair or  
 CC regeneration of tissue. AB280278 to AB280363 represent PCR primers which  
 CC are used in an example from the present invention for markers defining  
 CC cell populations

XX Sequence 24 BP; 7 A; 7 C; 6 G; 4 T; 0 U; 0 Other;

Query Match 10.9%; Score 24; DB 10; Length 24;  
 Best Local Similarity 100.0%; Pred.No. 0.089;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 108 CTCACCCCGCAGAGATTGTAGA 131  
 |||||  
 Db 1 CTCACCCCGCAGAGATTGTAGA 24

#### RESULT 2

AD007907  
 ID AD007907 standard; DNA; 23 BP.

XX AD007907;

XX 07-OCT-2004 (first entry)

XX Sense primer for RT-PCR analysis of PAX6 gene.

XX Antidiabetic; human hepatic cell; insulin; FH-B-TPN; type 1; diabetes;

XX pancreatic duodenal homeobox gene-1; Pdx 1; insulin-dependent; PCR;

XX primer; ss.

XX Unidentified.

XX WO2004061091-A2.

XX 22-JUL-2004.

XX 05-JAN-2004; 2004WO-GB000005.

XX 06-JAN-2003; 2003GB-00000208.

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX Bifrat S;

XX WPI; 2004-534380/51.

XX New modified human hepatic cell, which is insulin-producing, useful as a

XX medicament for treating insulin-dependent diabetes.

XX Example; Fig 9; 46pp; English.

XX The invention relates to a modified human hepatic cell, which is insulin-  
 CC producing. Further disclosed is an insulin-producing cell line designated  
 CC FH-B-TPN, a method of making an insulin-producing cell line, a method of  
 CC treatment of type 1 diabetes, a method of making insulin in-vitro, and a  
 CC vector for the expression of pancreatic duodenal homeobox gene-1 (Pdx1),

CC where the vector is a lentiviral vector. The modified human hepatic cell  
 CC is useful for treating insulin-dependent diabetes. Sequences given in  
 CC records for AD007888-AD007961 represent primers for the RT-PCR analysis  
 CC of RNA extracted from cultured cells, mature human hepatocytes, and human  
 CC pancreatic islets.

XX Sequence 23 BP; 7 A; 6 C; 7 G; 3 T; 0 U; 0 Other;

Query Match 9.5%; Score 21; DB 12; Length 23;  
 Best Local Similarity 100.0%; Pred.No. 2.8;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 CTCACAGCGGAGTGAAATCAGC 70  
 |||||  
 Db 3 CTCACAGCGGAGTGAAATCAGC 23

RESULT 3  
 AA016077  
 ID AA016077 standard; DNA; 38 BP.

XX AA016077;

XX 21-MAY-1998 (first entry)

XX PCR primer used to identify PAX6 mutations in mice.

XX Mutation; mutational screening; recessive; phenotypic alteration;

XX single strand conformation polymorphism; SSCP; PAX6 gene; aniridia;

XX PCR primer; amplify; ss.

XX Synthetic.

XX Mus sp.

XX WO9744485-A1.

XX 27-NOV-1997.

XX 16-MAY-1997; 97WO-GB001354.

XX 17-MAY-1996; 96GB-00010355.

XX (HEXA-) HEXAGEN TECHNOLOGY LTD.

XX Goodfellow PN;

XX WPI; 1998-018536/02.

XX Identification of mutation(s) in genes of interest - without prior

XX observation of phenotypic alteration in the mutated organism or cell.

XX Example 11; Page 58; 66pp; English.

XX PCR primers AA016059-76 were used to identify PAX6 mutations in mice

XX using the method of the invention. The method comprises testing a nucleic

XX acid sample from a mutated organism for a mutation in a gene of interest

XX without the prior observation of a phenotypic alteration in the mutated

XX organism resulting from the mutation. PAX6 mutations lead to a variety of

XX anterior segment malformations most commonly characterised by eye

XX development defects broadly described as aniridia. The disease is

XX dominant. A population of male mice were treated with ENVU to provide a

XX source of mutant PAX6 and a heterozygotic F1 generation produced.  
 CC Fluorescent single strand conformation polymorphism (SSCP) is utilised to  
 CC identify those members of the F1 population carrying PAX6 mutations. The  
 CC method provides mutational screening based on genomic and genetic  
 CC techniques rather than on phenotypic observation. The method identifies  
 CC and characterises genes via mutagenesis to identify genes encoding  
 CC products which may have therapeutic benefit. The method also identifies  
 CC the presence of mutations in a gene which do not rely solely upon prior  
 CC matching of a gene with a disease. Heterozygotic organisms can also be  
 CC screened to identify those carrying a mutation in a copy of a gene of  
 CC interest even though the gene may be recessive and therefore causes no  
 CC phenotypic alteration



AUTHORS Goodfellow,P.N.  
TITLE Methods for identifying a mutation in a gene of interest without a phenotypic guide  
JOURNAL Patent: US 5994075-A 149 30-NOV-1999;  
FEATURES Location/Qualifiers  
source 1..38  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 9.1%; Score 20; DB 6; Length 38;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 GTCACAGCGAGTGAATCAG 69  
Db 19 GTCACAGCGAGTGAATCAG 38  
|||||

RESULT 3  
A67658/c A67658 40 bp DNA linear PAT 05-MAY-1999  
LOCUS DEFINITION Sequence 78 from Patent WO9744485.  
ACCESSION A67658  
VERSION A67658.1 GI:4756521  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 40)  
AUTHORS Goodfellow,P.N.  
TITLE METHODS FOR IDENTIFYING A MUTATION IN A GENE OF INTEREST  
JOURNAL Patent: WO 9744485-A 78 27-NOV-1997;  
FEATURES HEXAGEN TECHNOLOGY LIMITED (GB)  
source 1..40  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"

ORIGIN

Query Match 9.1%; Score 20; DB 6; Length 40;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 160 GACATTCCCGAATCTGCA 179  
Db 40 GACATTCCCGAATCTGCA 21  
|||||

RESULT 4  
AR089868/c AR089868 40 bp DNA linear PAT 07-SEP-2000  
LOCUS DEFINITION Sequence 150 from patent US 5994075.  
ACCESSION AR089868  
VERSION AR089868.1 GI:10016623  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 40)  
TITLE Goodfellow,P.N.  
JOURNAL Methods for identifying a mutation in a gene of interest without a phenotypic guide  
FEATURES Patent: US 5994075-A 150 30-NOV-1999;  
source Location/Qualifiers  
1..40  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 9.1%; Score 20; DB 6; Length 40;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 160 GACATTCCCGAATCTGCA 179  
Db 40 GACATTCCCGAATCTGCA 21  
|||||

RESULT 5  
BD061402/c BD061402 21 bp DNA linear PAT 27-AUG-2002  
LOCUS DEFINITION Peptide having a function regulating transcription of gene.  
ACCESSION BD061402  
VERSION BD061402.1 GI:22607008  
KEYWORDS JP 2001292776-A/2.  
SOURCE JP 2001292776-A/2.  
ORGANISM synthetic construct  
REFERENCE other sequences; artificial sequences.  
1 (bases 1 to 21)  
AUTHORS Takagi,M., Shinji,H. and Ota,K.  
TITLE Peptide having a function regulating transcription of gene  
JOURNAL Patent: JP 2001292776-A 2 23-OCT-2001;  
COMMENT AGENCY OF IND SCIENCE & TECHNOL  
OS Artificial Sequence  
PN JP 2001292776-A/2  
PD 23-OCT-2001  
PF 11-APR-2000 JP 2000109760  
PI MASARU TAKAGI,HIDEAKI SHINJI,KEN OTA  
PC C12N15/09,C07K14/415,C12N5/10// (C12N15/09,C12R1:91),C12N15/00,  
PC C12N5/00,C12R1:91)  
PC (C12N15/00,C12R1:91)  
CC Description of Artificial Sequence: Synthetic primer DNA FH  
FEATURES Key  
source 1..21  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"

ORIGIN

Query Match 7.3%; Score 16; DB 6; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.6e+03;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 CCTCTCCCTTCTCTCA 48  
Db 18 CCTCTCCCTTCTCTCA 3  
|||||

RESULT 6  
AR036378 AR036378 34 bp DNA linear PAT 29-SEP-1999  
LOCUS DEFINITION Sequence 41 from patent US 5872105.  
ACCESSION AR036378  
VERSION AR036378.1 GI:5953046  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 34)  
TITLE Kool,E.T.  
JOURNAL Single-stranded circular oligonucleotides useful for drug delivery  
FEATURES Patent: US 5872105-A 41 16-FEB-1999;  
source Location/Qualifiers  
1..34  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 7.3%; Score 16; DB 6; Length 34;  
Best Local Similarity 100.0%; Pred. No. 2.6e+03;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 TTCCTTCCCTTCC 29  
|||||